

## Appendix I: Alignment of NP\_003204 with AAC50763 (SEQ ID NO: 7)

BLASTP 2.2.23+

Reference: Stephen F. Altschul, Thomas L. Madden, Alejandro A. Schaffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997), "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", *Nucleic Acids Res.* 25:3389-3402.

Reference for compositional score matrix adjustment: Stephen F. Altschul, John C. Wootton, E. Michael Gertz, Richa Agarwala, Aleksandr Morgulis, Alejandro A. Schaffer, and Yi-Kuo Yu (2005) "Protein database searches using compositionally adjusted substitution matrices", *FEBS J.* 272:5101-5109.

RID: Z8246KT01S

Database: All non-redundant GenBank CDS translations+PDB+SwissProt+PIR+PRF excluding environmental samples from WGS projects

11,049,075 sequences; 3,763,785,421 total letters  
Query= gi|4507427|ref|NP\_003204.1| TEA domain family member 4 [Homo sapiens]  
Length=434

## ALIGNMENTS

>gb|AAC50763.1| transcription factor RTEF-1 [Homo sapiens]  
Length=434

Score = 907 bits (2345), Expect = 0.0, Method: Compositional matrix adjust.  
Identities = 434/434 (100%), Positives = 434/434 (100%), Gaps = 0/434 (0%)

Query	1	MEGTAGTITSNEWSSPTSPGEGSTASGGSQALDKPIDNDGEGVWSPDIEQSFQEALAIYPP	60
Sbjct	1	MEGTAGTITSNEWSSPTSPGEGSTASGGSQALDKPIDNDGEGVWSPDIEQSFQEALAIYPP	60
Query	61	CGRRKIILSDEGKMYGRNELIARYIKLRTGKTRTRKQVSSHQVLARRKAREIQAKLKDQ	120
Sbjct	61	CGRRKIILSDEGKMYGRNELIARYIKLRTGKTRTRKQVSSHQVLARRKAREIQAKLKDQ	120
Query	121	AAKDKALQSMAMSSAQIISATAFHSSMRLARGPGRPAVSGFWQGALPGQAETSHDVKPF	180
Sbjct	121	AAKDKALQSMAMSSAQIISATAFHSSMRLARGPGRPAVSGFWQGALPGQAETSHDVKPF	180
Query	181	SQQTAYVQPPPLPLPGFESPAGPAPSPSAPPAPPPWQRRRSGSKLWMLEFSAFLEQQQDDPD	240
Sbjct	181	SQQTAYVQPPPLPLPGFESPAGPAPSPSAPPAPPPWQRRRSGSKLWMLEFSAFLEQQQDDPD	240
Query	241	TYNKHFLVHIGQSSPSYLRPYLEAVDIRIQYDKFPEKKGGLKDLFERGPSNAFFLVKFWA	300
Sbjct	241	TYNKHFLVHIGQSSPSYLRPYLEAVDIRIQYDKFPEKKGGLKDLFERGPSNAFFLVKFWA	300
Query	301	DLNTNIEDEGSSFFYGVSSQYESPENMIITCSTKVCSPGKQVVEKVEYARYENGHYSYR	360
Sbjct	301	DLNTNIEDEGSSFFYGVSSQYESPENMIITCSTKVCSPGKQVVEKVEYARYENGHYSYR	360
Query	361	IHRSPICEYMINFIHKLKHLPEKYMNSVLENFTILQVNVNRDQTETLLCIAYVFEVSAS	420

Sbjct	361	IHRSPICEYMINF I H K L K H L P E K Y M M N S V L E N F T I L Q V V T N R D T Q E T L L C I A Y V F E V S A S	
		IHRSPICEYMINF I H K L K H L P E K Y M M N S V L E N F T I L Q V V T N R D T Q E T L L C I A Y V F E V S A S	420
Query	421	EHGAQHHIYRLVKE	434
		EHGAQHHIYRLVKE	
Sbjct	421	EHGAQHHIYRLVKE	434